ARIZONA DEPARTMENT OF TRANSPORTATION

REPORT NUMBER: FHWA-AZ94-383

RHODES PROJECT: PHASE II (A)

Final Report

Prepared by:
Larry Head
Pitu Mirchandani
Systems and industrial Engineering Department
University of Arizona
Tucson, Arizona 85721

July 1994

Prepared for:
Arizona Department of Transportation
206 South 17th Avenue
Phoenix, Arizona 85007
in cooperation with
U S. Department of Transportation
Federal Highway Administration

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Arizona Department of Transportation or the Federal Highways Administration. This report does not constitute a standard, specification, or regulation. Trade or manufacturer's names which may appear herein are cited only because they are considered essential to the objectives of the report. The U.S. Government and the State of Arizona do not endorse products or manufacturers.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-AZ-94-383		

4. Title and Subtitle	5. Report Date
RHODES Project:Phase II (a)	July 1994
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Larry Head and Pitu Mirchandani	SIE Dept., University of Arizona

9. Performing Organization Name and Address	10. Work Unit No.
Systems and Industrial Engineering Department University of Arizona Tucson, AZ 85721	
	11. Contract or Grant No. HPR-PL-1(43)383

12. Sponsoring Agency Name and Address	13.Type of Report & Period Covered
ARIZONA DEPARTMENT OF TRANSPORTATION **206 S. 17TH AVENUE** **PHOENIX, ARIZONA 85007**	Final Report 5/92 -1/93
	14. Sponsoring Agency Code

15. Supplementary Notes

Prepared in cooperation with the U.S. Department of Transportation, Federal Highway Administration

16. Abstract

This report documents the work performed on the *RHODES* Project Phase II(a). This research effort was the continuation of the *RHODES* Project Phase I. Phase I explored concepts for models and algorithms for a real-time traffic-adaptive control systems for street networks, referred to as the *RHODES* System. Phase II(a) focused on further development of some of these algorithms and on performing some preliminary laboratory experiments with these algorithms using simulation models.

The control architecture of *RHODES* is based on a hierarchical decomposition of the overall traffic control problem. In an aggregate sense, there are three levels in the control hierarchy: *network load control*, *network flow (platoon) control*, and *intersection (vehicular) control*. *RHODES* architecture allows for a modular implementation of many of the subsystems within the control structure and the incorporation of IVHS technologies (e.g. new vehicle sensors) when they become available. In Phase II(a), the decision problems at each of the hierarchical levels were further analyzed and the decision model/algorithm at the intersection level was explicitly formulated, solved, and evaluated using simulation models.

17. Key Words	18. Distribution Statement	23. Registrant's Seal
Traffic Control Systems, intersection Control, Real-Time TrafficAdaptive Control, Simulation, Optimization.	Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161	

19. Security Classification	20. Security Classification	21. No. of Pages	22. Price
Unclassified	Unclassified	170	

METRIC (SI*) CONVERSION FACTORS

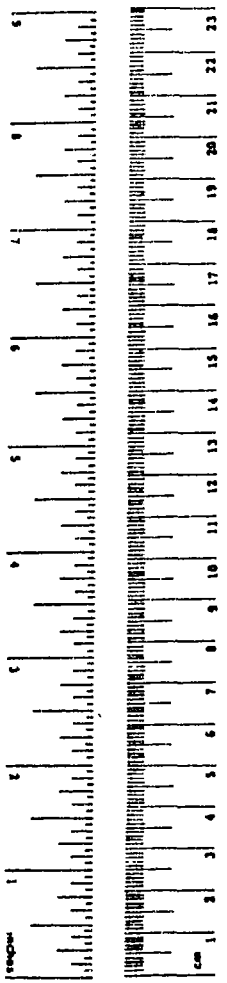
APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
in	inches	2.54	centimetres	cm
ft	feet	0.3048	metres	m
yd	yards	0.914	metres	m
mi	miles	1.61	kilometres	km
AREA				
in^2	square inches	645.2	centimetres squared	cm^2
ft^2	square feet	0.0929	metres squared	m^2
yd^2	square yards	0.836	metres squared	m^2
mi^2	square miles	2.59	kilometres squared	km^2
ac	acres	0.395	hectares	ha
MASS (weight)				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
VOLUME				
fl oz	fluid ounces	29.57	millilitres	mL
gal	gallons	3.785	litres	L
ft^3	cubic feet	0.0328	metres cubed	m^3
yd^3	cubic yards	0.0765	metres cubed	m^3

NOTE: Volumes greater than 1000 L shall be shown in m^3.

TEMPERATURE (exact)

°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

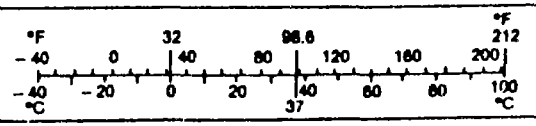


APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
mm	millimetres	0.039	inches	in
m	metres	3.28	feet	ft
m	metres	1.09	yards	yd
km	kilometres	0.621	miles	mi
AREA				
mm^2	millimetres squared	0.0016	square inches	in^2
m^2	metres squared	10.764	square feet	ft^2
km^2	kilometres squared	0.39	square miles	mi^2
ha	hectores (10 000 m^2)	2.53	acres	ac
MASS (weight)				
g	grams	0.0353	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams (1 000 kg)	1.103	short tons	T
VOLUME				
mL	millilitres	0.034	fluid ounces	fl oz
L	litres	0.264	gallons	gal
m^3	metres cubed	35.315	cubic feet	ft^3
m^3	metres cubed	1.308	cubic yards	yd^3

TEMPERATURE (exact)

°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



These factors conform to the requirement of FHWA Order 5190.1A.

* SI is the symbol for the International System of Measurements

PREFACE

This report documents the work performed on the *RHODES* Project Phase II(a). This research effort was the continuation of the *RHODES* Project Phase I. Both phases were funded by the Arizona Department of Transportation (ADOT) through the Pima Association of Governments (PAG). Essentially, Phase I explored concepts for models and algorithms for real-time traffic-adaptive control systems for street networks. Phase II(a) focused on further development of some of these algorithms and on performing some preliminary laboratory experiments with these algorithms using simulation models.

This report was written by the principal investigator, **Pitu B. Mirchandani**, and co-principal investigator, **Larry Head**, both of the Systems and Industrial Engineering (SIE) Department at the University of Arizona. It is based on the compilation of research efforts and results of various individuals who have been involved in the *RHODES* Project. In particular, the efforts of the following individuals are acknowledged:

Julia Higle	Associate Professor, SIE Department
Suvrajeet Sen	Associate Professor, SIE Department
Douglas Gettman	Undergraduate Assistant, SIE Department
Srikanth Nagarajan	Graduate Assistant, SIE Department
Ranjit Rebello	Graduate Assistant, SIE Department
Douglas Tarico	Graduate Assistant, SIE Department
Gregory Tomooka	Graduate Assistant, SIE Department
Michael Whalen	Graduate Assistant, SIE Department
Paolo Dell'Olmo	Visiting Scientist, SIE Department (from the Instituto di Analisi dei Sistemi ed Informatica, Rome).

In addition, the principal investigators wish to acknowledge the following individuals whose technical input, interactions with key investigators, and general support on the project has been invaluable:

Arizona Department of Transportation	
Harry A. Reed	Director, Transportation Planning, ADOT
Larry A. Scofield	Manager, Transportation Research, Arizona Transportation Research Center (ATRC), ADOT
Sarath Joshua	Senior Research Engineer, ATRC, ADOT
Louis Schmitt	Assistant County Manager, Maricopa County Transportation and Development Agency, (formerly with ADOT)
Pima Association of Governments	
Thomas Buick	Chief, Transportation Planning Division, Maricopa County Department of Transportation, (formerly with PAG)
David Wolfson	Senior Transportation Planner, PAG
City of Tucson	
Benny Young	Director of Transportation, City of Tucson
Richard Nassi	Traffic Engineer, City of Tucson.
Dennis Sheppard	Assistant Traffic Engineer, City of Tucson.

Comments and support of **Harvey Friedson** and **James Decker** of the Traffic Engineering Department of the City of Tempe, AZ, are also gratefully acknowledged.

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views of the Arizona Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification or regulation.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF TABLES

1. INTRODUCTION

Over a period of 15 months, since June 1, 1991, the Arizona Department of Transportation supported the R&D efforts on the development of the RHODES street traffic control system within the Department of Systems and Industrial Engineering at the University of Arizona. PHASE I and PHASE II(a) of this effort have been completed. During these phases the University of Arizona has worked closely with the City of Tucson and the Pima Association of Governments (PAG) in the development of the RHODES concept, some preliminary algorithms, and a simulation model.

PHASE I of the RHODES project consisted of the following tasks:

Task 1(a): Develop RHODES concept
Task 1(b): Develop analysis/simulation tools
Task 1(c): Select demonstration test grid
Task 1(d); Hold traffic control workshop

Task 2(a): Refine RHODES concepts
Task 2(b): Investigate flow optimization models
Task 2(c): Investigate intersection dispatching schemes
Task 2(d): Coordinate modeling efforts.

Tasks 1(a) and 2(a) concentrated on developing a technically sound concept for real-time traffic adaptive control and identifying the key research problems that need to be solved. Task 1(b) consisted of specifying the requirements for a simulation model for demonstrating, testing and evaluating real-time control. It was decided, at least for the short term, that modification of the TRAF-NETSIM model would provide a suitable simulation environment. In the longer term, more advanced simulation models that allow dynamic vehicle routing and have the ability to assess ATIS and other IVHS technologies would be more appropriate. To this end, an investigation of object oriented traffic simulation has been initiated.

Task 1(c) addressed the long-term project goal of implementing RHODES for the Tucson street network. A potential test grid has been selected. The test grid offers several interesting traffic characteristics, such as having a variety of traffic volumes, and a mix of residential and commercial zones. Early selection of the test grid provides a source of

real-world data for the traffic simulations and a measure against which to validate the developed simulation model.

Task 1(d) provided a forum of noted experts on real-time traffic control to discuss research issues and comment on the RHODES concept. The workshop was very valuable to the research team. It led to the refinement of the RHODES concept and identified several new key issues.

Tasks 2(b) and (c) focused on the investigation and development of some preliminary algorithms for intersection and network flow control. An algorithm was developed, called COP, based on a dynamic programming formulation of the intersection control problem. The COP algorithm provides the necessary planning horizon, approximately 5 minutes, for integration with network flow control methods.

In addition to these tasks, a major goal of PHASE I was the development of a proposal to FHWA on the design of a real-time traffic-adaptive signal control system. The RHODES team led a strong consortium, that included JHK, SRI, TASC, RPI, and Hughes, and submitted a consortium proposal to FHWA in January 1992. The proposal was not selected for funding; the contract was awarded to Farradyne Systems in June 1992. However, the RHODES team plans to respond to an anticipated FHWA-RFP that will call for alternative prototype developments.

Phase II(a) is concentrated on (1) the development of some RHODES component models and algorithms and (2) a demonstration of these algorithms, using the modified TRAF-NETSIM simulation model. PHASE II(a) consisted of three task:

Task A: Develop algorithms for network loading and control
Task B: Demonstrate controller interface and network control concepts
Task C: Reporting and planning

Task A addressed the investigation and development of algorithms at several levels of the hierarchy as identified in PHASE I. The purpose of Task B is to demonstrate the proof of concept that the RHODES approach can be implemented using existing controller technology.

The research progress on the RHODES project has been significant. A simulation model has been developed for testing and demonstrating real-time traffic control algorithms, and

several algorithms have been developed. Concurrently with the further development of the RHODES system for street network control, it is now appropriate to extend the RHODES concept for developing a traffic control system for an integrated freeway/street network. This final report contains the detailed results of the PHASE II(a) effort.

The RHODES concept is depicted in Figure 1. At the highest level of RHODES is the *"dynamic network loading"* model that captures the slow-varying characteristics of traffic. These characteristics pertain to the network geometry (available routes including road closures, construction, etc.) and the typical route selection of travelers. Based on the slow-varying characteristics of the network traffic loads, estimates of the load on each particular link, in terms of vehicles per hour, can be calculated. These load estimates then allow RHODES to allocate "green time" for each different demand pattern and each phase (North-South through movement, North-South left turn, East-West left turn, and so on). These decisions are made at the middle level of the hierarchy, referred to as *"network flow control"*. Traffic flow characteristics at this level are measured in terms of platoons of vehicles and their speeds. Given the approximate green times, the *"intersection control"* at the third level selects the appropriate phase change epochs based on observed and predicted arrivals, of individual vehicles, at each intersection. The RHODES architecture is modular; it allows the accommodation of new modeling methodologies and new technologies as they are developed.

A significant difference between RHODES and other "real-time" traffic control systems is that RHODES is being designed to accommodate real-time measurements of traffic and to become an integral component of IVHS. For example, integration of *Advanced Traveler Information Services* within IVHS will result in (1) improved prediction and estimation of network loads, (2) will allow the ATMS system to provide drivers with real-time information about traffic conditions, and (3) advise the travelers of alternate routes. Priority and accommodation of public and private transit, emergency vehicles, and commercial vehicles, can be easily integrated into the decision-making structure of RHODES.

At the highest *network loading* level of the hierarchy we envision the decision time horizons to be in hours, days and weeks. This model allows for integration of historical

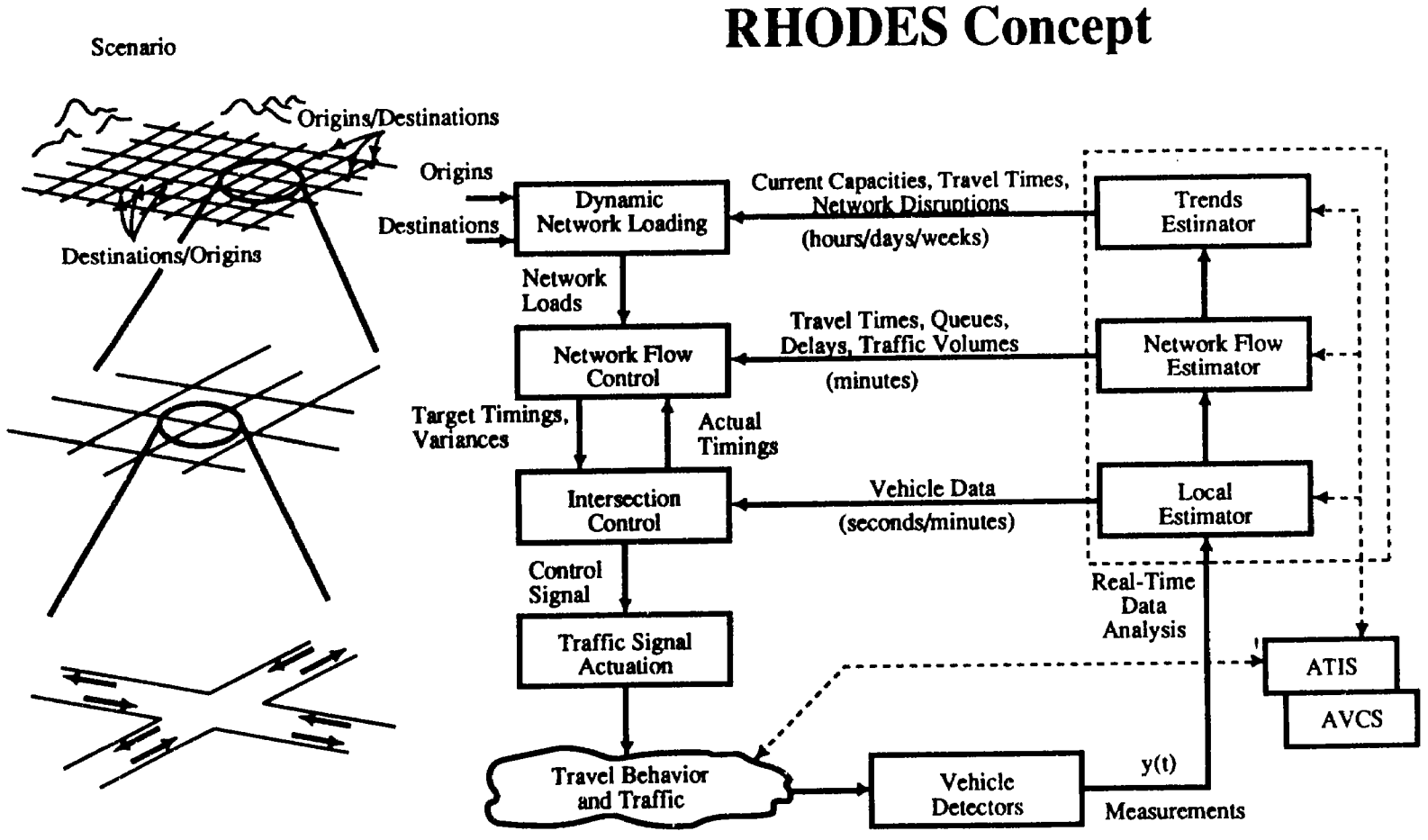


Figure 1. The Rhodes Hierarchical Control System Architecture.

data (a priori information), observed traffic flows (posterior information) and potential[1] ATIS information about IVHS suggested routes and traffic conditions (congestion, accidents and other network events) to allow prediction of near future loads and hence exercise real-time proactive traffic control. The next level of the hierarchy utilizes the predicted and estimated network loads to control traffic on a network wide basis. At this level the *network flow controller* will integrate the network load information with observations of actual volumes and flow profiles to select appropriate phase sequences and phase lengths as well as the allowable variances to accommodate for the stochastic nature of traffic flow on the network level. These timing decisions will be passed to the *intersection controller* where decisions to shorten or extend the current phase will be made (in a decentralized distributed fashion) based on actual observations of the current traffic arrival pattern at each intersection. The lowest level of the hierarchy, referred to as *traffic signal actuation*, is responsible for implementation of the intersection controller decision on the signal control hardware.

[1] The scope of this effort does not include development of an ATIS system. It does, however, include the consideration of potential information available from an ATIS in the design of a proactive traffic control system.

2. NETWORK LOADING

2.1 A Statistical Network Loading Model

In this section, a method that uses historical data to estimate network loads is described. The method is similar in spirit to the dynamic Bayes procedure described by Higle and Nagarajan (1992), although it has been adapted to context of network load estimation. The method is an empirical procedure where the amount of data used to obtain load estimates is determined by the quality/accuracy of the estimates being produced. Higle and Nagarajan show that the procedure is well suited to identifying and reacting to changes in the underlying traffic trends, specifically turning flow probabilities. Thus, it is believed that this method will also be well suited to estimating network loads.

2.1.1 Empirical Bayes Estimates

The primary objective is to estimate the number of vehicles traveling on a particular link during a particular interval of time on a particular calendar day. Let $N_{ij}(t,d)$ be the number of vehicles traveling on link (i,j) during time period t on calendar day d, as observed using vehicle detectors. Assume that $N_{ij}(t,d)$ has a Poisson distribution with a mean $\lambda_{ij}(t,d)$. There are several points implicit in this simple assumption.

First, note that the average vehicle load on link (i,j), denoted by $\lambda_{ij}(t,d)$ need not be presumed constant over time or over calendar day. This rate typically varies by "time of day" and "day of week". Figure 2 depicts the time varying vehicular flow on a particular day of the week for various calendar days. It is assumed that $\lambda_{ij}(t,d)$ for each day d is for a collection of calendar days that have essentially the same characteristics. Second, the interval of time, Δt, associated with the estimation/prediction task need not be held constant throughout the day, but may also vary by time of day. Thus, there is sufficient flexibility in the estimation/prediction method to allow for longer time intervals during low use periods, and shorter time intervals during high use periods. Third, as discussed before, the vehicular rate, $N_{ij}(t,d)$, is generally assumed constant for a particular time period on a particular day. Finally, note that $\lambda_{ij}(t,d)$, the average vehicle loads, are the quantities is to be estimated.

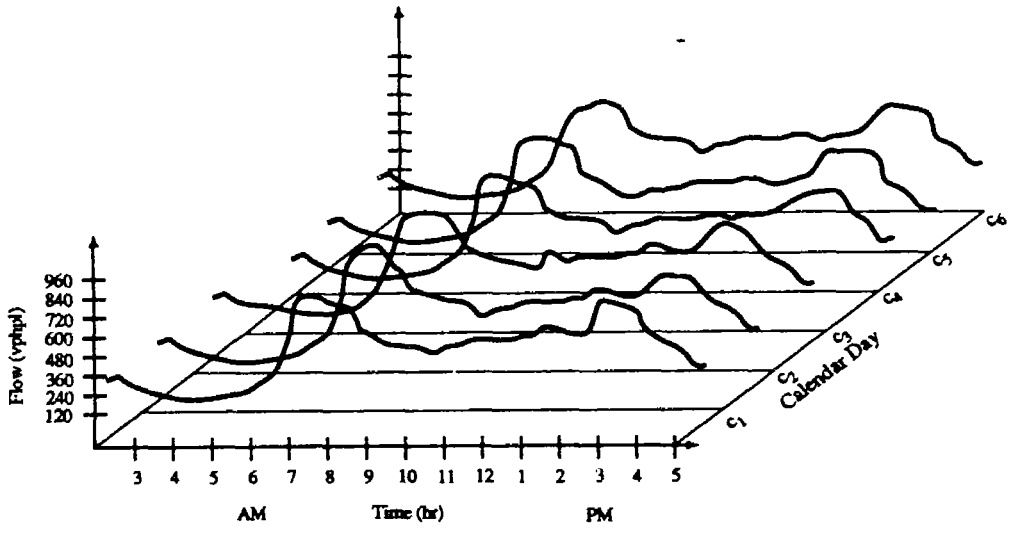


Figure 2. Typical day-of-week loads on link [i, j] over time of day and on different calendar days.

Since the mean, $\lambda_{ij}(t,d)$, of the Poisson distribution is unknown, a Bayesian viewpoint is adopted and it is modeled as a random variable. For the purposes of mathematical convenience and computational ease, assume that $\lambda_{ij}(t,d)$ has a gamma distribution with parameters α_0 and β_0. As data is collected the parameters of the gamma distribution will be updated to describe $\lambda_{ij}(t,d)$. The mean of the resulting distribution will be used as a point estimator for $\lambda_{ij}(t,d)$. That is, if $\lambda_{ij}(t,d) \sim \text{Gamma}(\alpha_0,\beta_0)$, then its mean

$$\hat{\lambda}_{ij}(t,d) = \frac{\alpha_0}{\beta_0} \tag{1}$$

will be used as the point estimate of the vehicle flow rate.

The estimation procedure evolves over time in a manner that follows readily from well known properties of the gamma and Poisson distributions. Specifically the key property is:

> If $\{N_k\}_{k=1}^{\delta}$ are independent and identically distributed observations of a random variable whose conditional distribution given μ is Poisson(with mean μ), and μ has a gamma distribution with parameters α_0 and β_0, then the conditional

distribution of μ, given the observations $\{N_k\}_{k=1}^{\delta}$ is a gamma distribution with parameters α_δ and β_δ, where

$$\alpha_\delta = \alpha_0 + \sum_{k=1}^{\delta} N_k \text{ and } \beta_\delta = \beta_0 + \delta, \tag{2}$$

(see, for example (DeGroot, 1977), chapter 11).

When translated to the context of traffic flow estimation, this result leads to a simple recursive procedure for estimating traffic flows over time. Here, μ corresponds to $\lambda_{ij}(t,d)$. Thus, given initial values of the parameters of the gamma distribution, α_0 and β_0, these values are updated to reflect the observations $N_{ij}(t,l), l = 1,\ldots,d$ as follows:

$$\begin{aligned} \alpha_d &= \alpha_0 + \sum_{l=1}^{d} N_{ij}(t,l), \\ \beta_d &= \beta_0 + d. \end{aligned} \tag{3}$$

This may be accomplished recursively as:

$$\begin{aligned} \alpha_d &= \alpha_{d-1} + N_{ij}(t,d-1), \\ \beta_d &= \beta_{d-1} + 1. \end{aligned} \tag{4}$$

Given α_d and β_d the predicted load on the forthcoming day is given by

$$\hat{\lambda}_{ij}(t,d+1) = \alpha_d/\beta_d. \tag{5}$$

To ensure that the resulting estimator, α_d/β_d, is capable of responding to changes in the underlying trends and responds adaptively to the quality of the estimates being produced, it is necessary to be able to use different amounts of data for obtaining α_d and β_d. That is, when the estimate is "good", additional data should be included so that estimates with lower error variances will result. However, when the estimate appears to be persistently poor, less, but newer, data should be used so that the estimators will be more responsive to apparent changes in the underlying trends. In the next section an adaptive method for determining the amount of data used in the estimation process is discussed.

2.1.2 Dynamic Bayes Estimation

The estimators obtained using the updated parameters specified in (3) and (4) will be most accurate when the underlying flow rate, $\lambda_{ij}(t,d)$, is constant over calendar day, d. However, even when the data is normalized for time of day and calendar day cycles, there are still likely to be changes in the flow rate (e.g., seasonal tendencies, construction obstructions, special events, etc.). In this case, it is necessary to allow the estimators to respond dynamically to the errors observed. This can be accomplished using the dynamic Bayes estimates proposed by Higle and Nagarajan (1992), adapted to the context of network load estimation.

Note first that the estimate of the anticipate flow for calendar day d is based on the observed flows in the previous calendar days. Thus, if α_d and β_d denote the parameters of the gamma distribution used to describe $\lambda_{ij}(t,d+1)$ after having observed $\left\{N_{ij}(t,l)\right\}_{l=1}^{d}$, then the point estimate of $\lambda_{ij}(t,d+1)$ is given by

$$\hat{\lambda}_{ij}(t,d+1)=\frac{\alpha_d}{\beta_d}. \tag{6}$$

The quality of this estimate depends on the extent to which it appears to be approximately equal to the observed flow in that period, $N_{ij}(t,d+1)$. If $N_{ij}(t,d+1)$ is consistent with $\hat{\lambda}_{ij}(t,d+1)$, then the vehicle flow rate over time appears to be stable enough to allow the simple update procedure described in (4). If $N_{ij}(t,d+1)$ is inconsistent with $\hat{\lambda}_{ij}(t,d+1)$, then steps must be taken to allow subsequent estimates to adapt to a potential change in the underlying trend. One approach is to discard the observations used early in the estimation process, so that the more recent observations influence the estimate more significantly.

To determine whether or not the observation is inconsistent with the estimate, it is necessary to obtain probabilistic statements from the Poisson distribution. Let an error probability ε, $0<\varepsilon<1$, be given and let quantities $\overline{N}$ and $\underline{N}$ be defined so that if $N_{ij}(t,d) \sim \text{Poisson}(\hat{\lambda}_{ij}(t,d))$ then

$$P\left\{N_{ij}(t,d) \le \underline{N}\right\} = \frac{\varepsilon}{2},$$
$$P\left\{N_{ij}(t,d) \ge \overline{N}\right\} = \frac{\varepsilon}{2}. \tag{7}$$

Therefore, under the hypothesis that $N_{ij}(t,d) \sim \text{Poisson}(\hat{\lambda}_{ij}(t,d))$,

$$P\left\{\underline{N} \le N_{ij}(t,d) \le \overline{N}\right\} = 1-\varepsilon. \tag{8}$$

In standard statistical tests of hypothesis, $N_{ij}(t,d)$ is said to be consistent with $\hat{\lambda}_{ij}(t,d)$ if $\underline{N} \le N_{ij}(t,d) \le \overline{N}$, and inconsistent otherwise. The inconsistency is said to be persistent if there is at least one inconsistency in the previous γ days, where γ is a specified estimation threshold. In this dynamic Bayes approach, α_d and β_d are determined in one of three ways, depending on the detection and persistence of an inconsistency.

Exhibit A outlines the dynamic Bayesian network loading algorithm. Note that at the start of this procedure, α_{d-1} , β_{d-1}, and therefore $\hat{\lambda}_{ij}(t,d)$ are available from the previous estimate. Parameters ε , δ and γ are chosen (by the designer/user of the algorithm) to facilitate the responsiveness of the algorithm to real-time data. As discussed above, ε and γ are used to identify inconsistancy in the estimates being produced, and δ denotes the time-window of the data used for the estimation. $\delta_{\min}$ and $\delta_{\max}$ are the minimum and maximum allowable sizes for the time window δ . The parameter δ is updated within the algorithm, depending on the quality of the estimates being produced. In steps 3 and 4, estimated α_d and β_d are derived using observed data $\left\{N_{ij}(t,l)\right\}_{l=d-d+1}^{d}$ during the current time window.

A brief discussion of Step 2 of this procedure is in order. Entering Step 2, $\hat{\lambda}_{ij}(t,d)$ has been determined from $\left\{N_{ij}(t,l)\right\}_{l=d-\delta}^{d-1}$ and is compared to the observed value $N_{ij}(t,d)$ to determine if δ should be adjusted. Note that δ is increased in Step 2 (a), decreased in Step 2 (b), but at all times $\delta_{\min} \le \delta \le \delta_{\max}$. These limits are intended to prevent the use of "too many" or "too few" observations in the computation of $\hat{\lambda}_{ij}(t,d)$ and can be specified at the user's discretion. In step 2 (a), the estimate is consistent with the observation, so δ is increased. In step 2 (b), the persistance of the inconsistancy is tested. If it is determined that the inconsistancy is persistant, δ may be decreased. Each time it is decreased, the oldest observation is discarded and a new estimate of $\hat{\lambda}_{ij}(t,d)$ is computed using the most recent observations (excluding, of course, $N_{ij}(t,d)$). The reduction of δ

Exhibit A. Dynamic Bayesian Network Loading Algorithm

Procedure: Compute $\hat{\lambda}_{ij}(t,d+1)$

Step 0:. Given estimates α_{d-1}, β_{d-1}, $\hat{\lambda}_{ij}(t,d)$, parameters γ, $0<\varepsilon<1$, δ, $\delta_{\min}$, $\delta_{\max}$ and the observations $\left\{N_{ij}(t,l)\right\}_{l=d-\delta+1}^{d}$ for time period t and caledar day d,

Step 1: Define $\overline{N}$ and $\underline{N}$ according to (7).

Step 2: Determine δ, the amount of data to be used in computing α_d and β_d:

(a) If $\underline{N} \le N_{ij}(t,d) \le \overline{N}$

then

$$\delta = \text{Min}(\delta+1, \delta_{\max}),$$

(b) else if $N_{ij}(t,l) \notin \left[\underline{N}, \overline{N}\right]$ for at least one observation $l \in [d-\gamma, d-1]$,

then while $N_{ij}(t,d) \notin \left[\underline{N}, \overline{N}\right]$ and $\delta > \delta_{\min}$

$$\begin{aligned}
\delta &\leftarrow \delta - 1 \\
\alpha_{d-1} &\leftarrow \alpha_{d-1} - N_{ij}(t, d-\delta) \\
\beta_{d-1} &\leftarrow \beta_{d-1} - 1 \\
\hat{\lambda}_{ij}(t,d) &\leftarrow \frac{\alpha_{d-1}}{\beta_{d-1}} \\
\alpha_0 &\leftarrow \alpha_{d-\delta} \\
\beta_0 &\leftarrow \beta_{d-\delta}
\end{aligned}$$

define $\underline{N}$, $\overline{N}$ according to (7) using $\hat{\lambda}_{ij}(t,d)$ and repeat **Step 2 (b)**

(c) else δ remains unchanged.

Step 3: Compute

$$\begin{aligned}
\alpha_d &\leftarrow \alpha_0 + \sum_{l=d-\delta+1}^{d} N_{ij}(t,l), \\
\beta_d &\leftarrow \beta_0 + \delta.
\end{aligned}$$

Step 4: Compute

$$\hat{\lambda}_{ij}(t,d+1) \leftarrow \frac{\alpha_d}{\beta_d}.$$

terminates when either $N_{ij}(t,d)$ becomes consistent with the recomputed estimate, $\hat{\lambda}_{ij}(t,d)$ or when δ reaches its lower limit, δ_{min}. When the algorithm process is in step 2 (c), inconsistency has been detected, but it is too early to tell if it is persistent. In this case, δ is neither increased nor decreased. Once δ has been set, $\hat{\lambda}_{ij}(t,d+1)$ is computed using the δ most recent observations, including $N_{ij}(t,d)$. By monitoring the quality of the estimates produced and adaptively responding to errors when they are detected, the estimated load for the forthcoming time period, $\hat{\lambda}_{ij}(t,d+1)$, should more closely approximate the load that will be observed , $N_{ij}(t,d+1)$.

2.2 A Network Loading Example

To demonstrate both the statistical network loading model and the capacity allocation model discussed above, a small traffic network was simulated using the modified TRAF-NETSIM model. Figure 3 shows the layout of the traffic network. This network was selected because it contains a long arterial (Campbell Avenue) near the University of Arizona football stadium[1]. The primary nodes of interest, those that will be used for testing control algorithms, are numbers 335, 369, 401 and 483. The remainder of the nodes are included to provide realistic traffic flows, i.e. platoons and non-uniform arrivals, into the controlled area. The location of vehicle detectors in the simulation model is consistent with the existing detector locations in the actual network. For the purposes of this example, the conditions on the network were simulated between 11 AM and 1 PM, a period of moderate to heavy usage.

To demonstrate the statistical network loading algorithm, the dynamic Bayes algorithm, detectors on each major links of the network are included in the estimation. For the purposes of presentation in this paper, the results from a single detector will be discussed in detail. This detector is located 130 feet north of the intersection of Speedway Blvd. and Campbell Ave (intersection 335). It includes all vehicles in all three lanes that approach the intersection. To represent both time periods and calendar days several runs of the simulation model were made. Each run utilized a unique random number with all other parameters (source input rates, turning probabilities, and signal timing parameters) held constant.

[1]This network selection is intended to allow the RHODES team to be prepared for the FHWA Real-time Traffic Adaptive Signal Control RFP for alternative algorithms due to be announced in 1993. This type of network/arterial will be the basis for the testing and performance competition for real-time traffic-adaptive control algorithms.

Figure 4 shows the observed number of vehicles for the eight time periods of fifteen minutes each over thirty days (simulation runs). From these observations it can be seen that the number of vehicles crossing the detector is essentially constant over the days and not constant over time. To test the estimation procedure the initial estimates, α_0 and β_0, were selected so that the estimate would beinitially be inaccurate and the responsiveness of the algorithm could be validated. Figure 5 shows the estimated loads over time and day. From the Figure it appears that the estimator can overcome the large initial estimate error and closely estimate the loads.

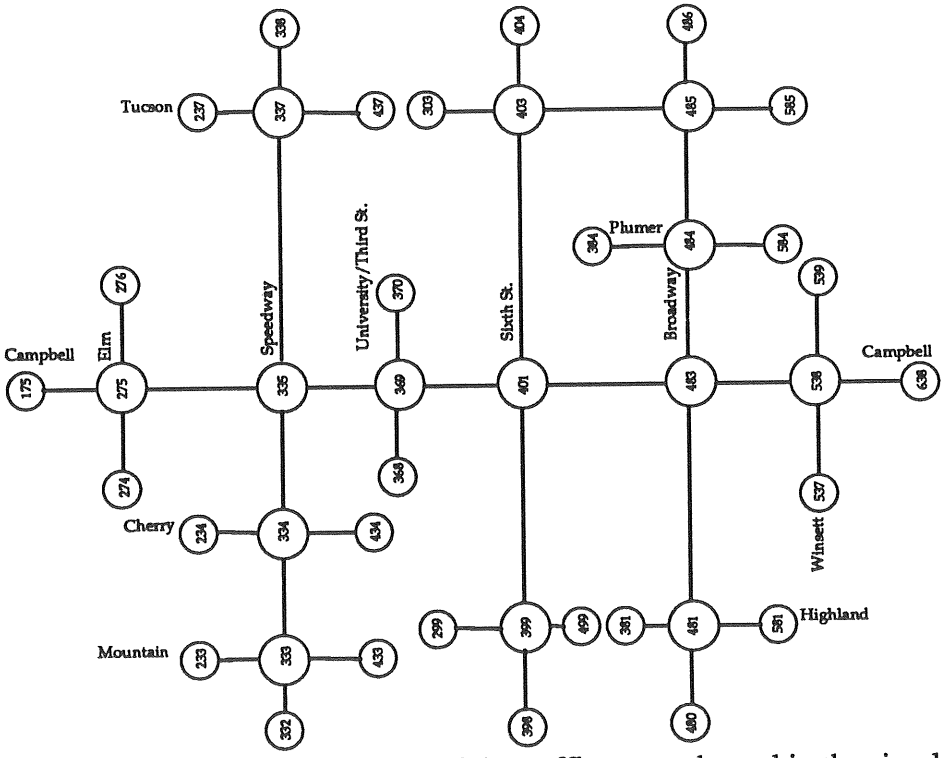


Figure 3. Topological layout of the traffic network used in the simulation studies.

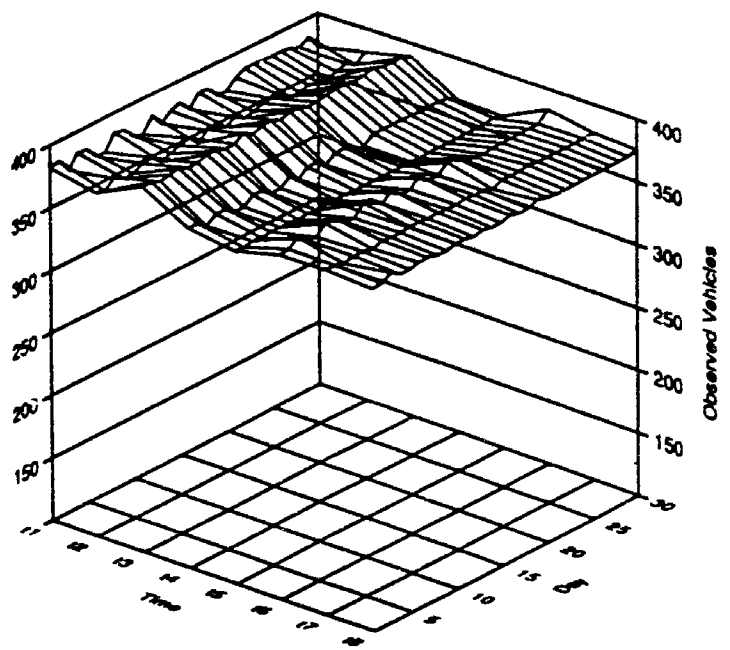


Figure 4. Observed number of vehicles for eight time periods of 15 minutes each over thirty days (simulation runs)

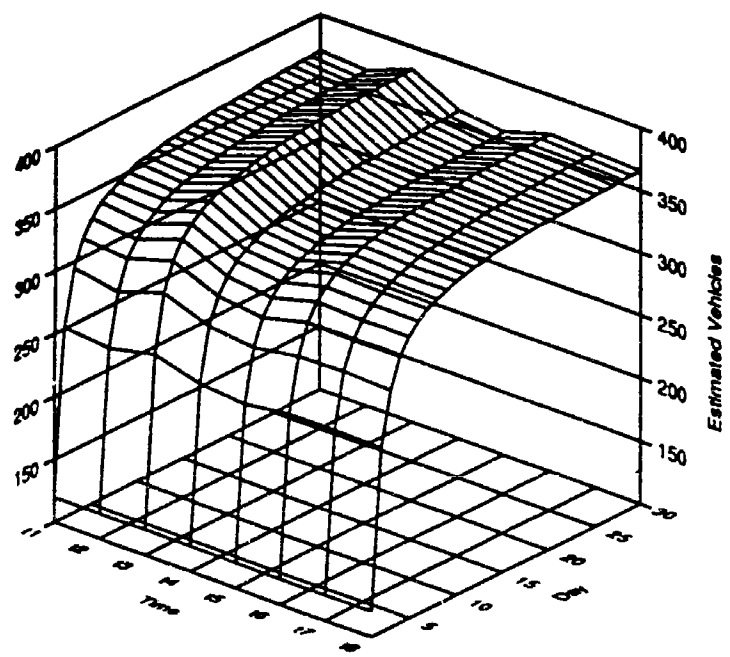


Figure 5. Estimated loads over time and day.

The performance of the method can be more closely studies by considering either a single time period or a single day. Figures 6, 7 and 8 show the observed and estimated loads over time on calendar days 0, 5 and 30, respectively. The large initial estimate error, approximately 100% on day 0, is reduced to less than 10% afer only five days and less than 1% ofter 30 days. Figure 9 shows the observed and estimted loads over calendar day at a single time period, t_3. This figure demonstrates the ability of the method to correctly estimate the load.

The statistical network loading model presented here is useful for estimating the expected link volumes based on existing loop detector data. It is important to note that this model is not based on known, or approximated, origin-destination data and hence is not an equilibrium or assignment model. This model does address the need for a statistical method of estimating link volumes based on loop detector data that will allow for the statistical classification of anomalies such as non-recurrent congestion due to events such as accidents. In these cases, alternative historical data sets can be used for the prediction purpose. Based on this statistical foundation this model can be extended to include equilibrium or assignment data, as well as other information that will be available through the deployment of IVHS.

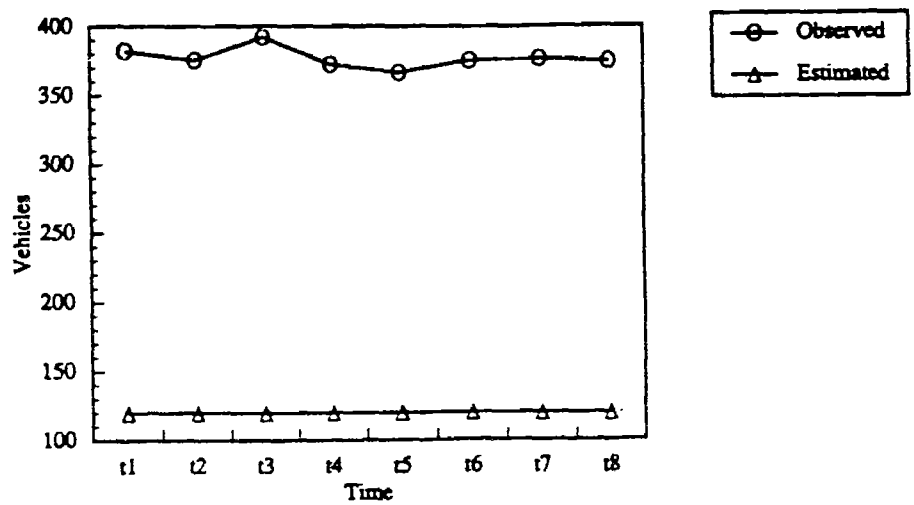


Figure 6. Observed and estimated load over time on calendar day 0.

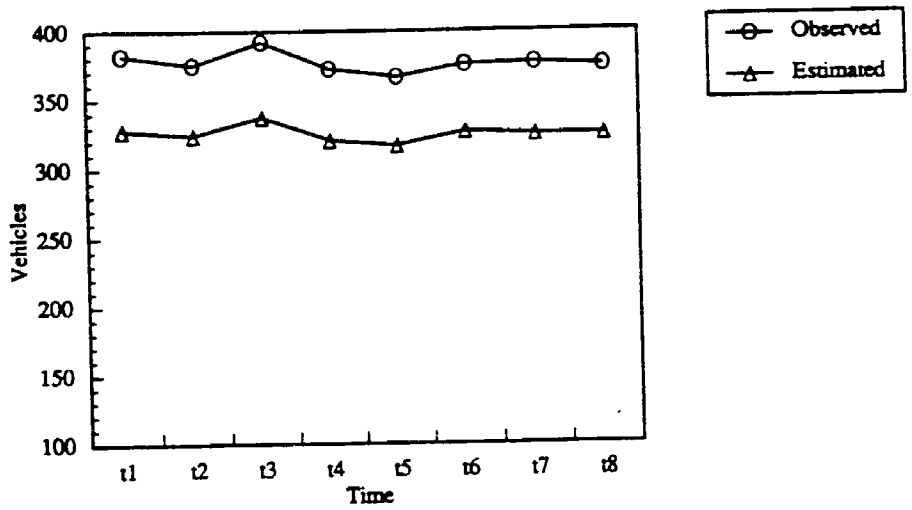


Figure 7. Observed and estimated load over time on calendar day 5.

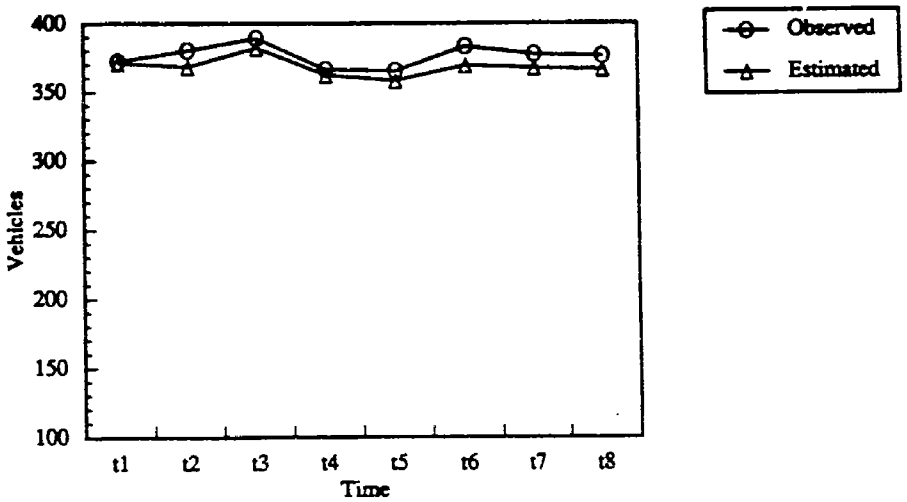


Figure 8. Observed and estimated load over time on calendar day 30.

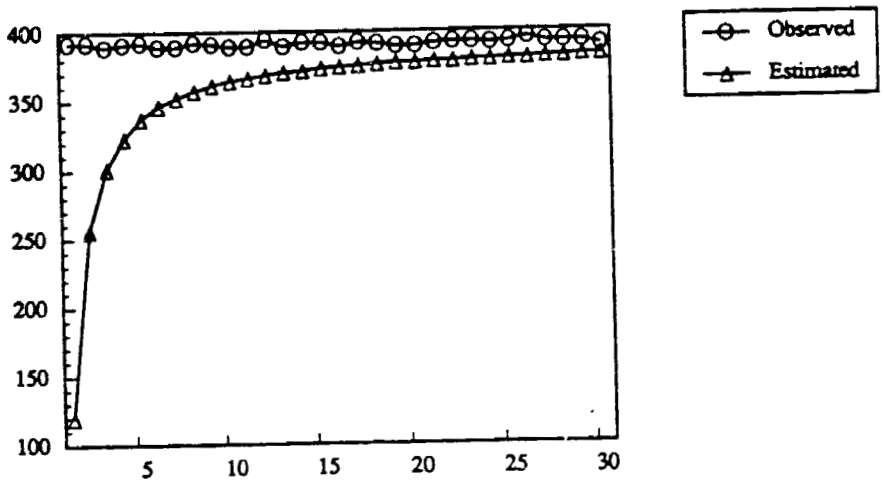


Figure 9. Observed and estimated load over calendar day at a fixed time.

3. CAPACITY ALLOCATION

3.1 The Capacity Allocation Model

The network loading model provides estimates of the expected link loads on the network. These estimates are used by the capacity allocation model to determine the fraction of time that should be allocated to each phase in order to satisfy the network demand. At this level of the hierarchy, a uniform, fluid flow viewpoint of traffic is assumed. The solution to the capacity allocation problem does not consider the flow of individual vehicles or platoons between signalized intersections. It establishes general fractions of time that must be allocated to different phases to satisfy the average demand over extended periods of time. These fractions serve as constraints to the network coordination model and the intersection scheduler.

Let v_i denote the demand (arrival rate) for movement i at some intersection. This demand can be derived from the predicted loads generated by the network loading model and estimated turning probabilities[1]. The quantity $\hat{\lambda}_{ij}(t,d+1)$ represents the estimated vehicular load on link (i,j) during time period t and calendar day d. If p_{ij}^m denotes the probability of a vehicle on link (i,j) demanding movement m then

$$v_m = p_{ij}^m \hat{\lambda}_{ij}(t,d+1) \tag{9}$$

is the estimated demand for movement m.

Figure 10 shows the standard labels of 8 possible movements at an intersection. Let $\phi_{ij} = \{i,j\}$ denote the signal phase where movements i and j are allowed. For the purpose of this development, assume that the only possible phases at this intersection are $\phi_{26}, \phi_{15}, \phi_{48}$ and ϕ_{37} (as shown in Figure 10). Let x_{26}, x_{15}, x_{48}, and x_{37} denote the fractions of the intersection capacity (green time) allocated to each phase. Then, assuming a uniform arrival rate, the delay (Hurdle, 1984) (uniform delay per vehicle) associated with phase ϕ is

[1] Here it is assumed that the turning probabilities are known. Estimation of these turning probabilities has been address by Higle and Nagaragan (1992) for the case of a fully instrumented intersection. Their approach has still to be adapted to a partially instrumented intersection.

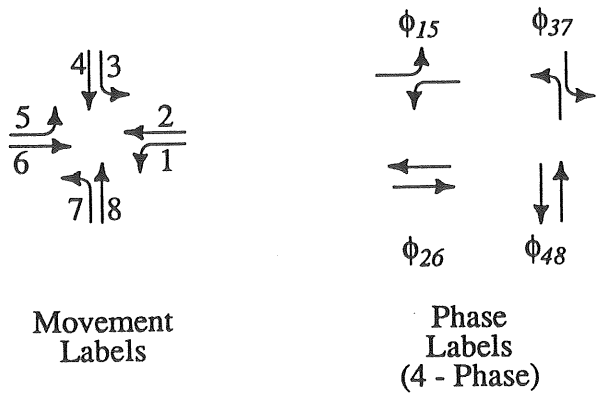


Figure 10. Standard labels of 8 possible movements and the associated 4-phases.

$$D(x_\phi) = \frac{(1-x_\phi)^2}{2} \sum_{m \in \phi} \frac{1}{(1-\frac{v_m}{s_m})} \tag{10}$$

where s_m denotes the saturation flow rate associated with movement m. Note that the saturation flow rate must be selected to reflect the appropriate number of lanes and other traffic considerations (grade, lane width, etc.) associated with each movement.

Given these definitions of delay, arrival rates and saturation flow rates, the capacity allocation problem can be stated as

$$\begin{aligned} &\text{Minimize } D(x) = \sum_{\text{all } \phi} D(x_\phi) \\ &\text{subject to} \\ &\qquad \sum_{\text{all } \phi} x_\phi = 1 \\ &\qquad x_\phi \geq 0, \quad \text{all } \phi. \end{aligned} \tag{11}$$

For the 4-phase case depicted in Figure 10 the capacity allocation problem can be stated as

$$\text{Minimize } D(x) = \frac{(1-x_{15})^2}{2}\left[\frac{1}{(1-\frac{v_1}{s_1})}+\frac{1}{(1-\frac{v_5}{s_5})}\right]+\frac{(1-x_{26})^2}{2}\left[\frac{1}{(1-\frac{v_2}{s_2})}+\frac{1}{(1-\frac{v_6}{s_6})}\right]$$
$$+\frac{(1-x_{37})^2}{2}\left[\frac{1}{(1-\frac{v_3}{s_3})}+\frac{1}{(1-\frac{v_7}{s_7})}\right]+\frac{(1-x_{48})^2}{2}\left[\frac{1}{(1-\frac{v_4}{s_8})}+\frac{1}{(1-\frac{v_8}{s_8})}\right] \tag{12}$$

subject to

$$x_{15}+x_{26}+x_{37}+x_{48}=1,$$
$$x_{15}\geq 0, \quad x_{26}\geq 0,$$
$$x_{37}\geq 0, \quad x_{48}\geq 0.$$

The capacity allocation problem, as stated here, has a quadratic objective function with a single linear equality constraint and the usual non-negativity (inequality) constraints on the decision variables. This form of a quadratic mathematical programming problem can be solved using the quadratic structure of the objective function and an *active set method* to manage the inequality constraints.

The capacity allocation problem (12) can be written in the form

$$\text{Minimize } \frac{1}{2}x'Qx+c'x+K$$
$$\text{subject to} \tag{13}$$
$$e'x=1,$$
$$x_i\geq 0, \quad i=1,\ldots N,$$

where $e'=(1 \quad \cdots \quad 1)$ and N is the number of allowable phases. Given this formulation the only mathematical condition necessary for a quadratic programming active set method to be applicable is that the matrix Q be positive definite.

The capacity allocation algorithm is outlined in Exhibit B. It is assumed that an initial starting solution is given (step 0). It is important that the initial solution satisfy the equality constraint $e'x^0=1$. A good candidate is

$$x_i^0=\frac{\sum_{m\in\phi_i} v_m}{\sum_{\text{all }\phi} v_m}. \tag{14}$$

Exhibit B. Capacity Allocation Algorithm.

Procedure: Solve Quadratic Program Using Active Set

Step 0:. Assume x^0 , a starting solution, is given and feasible.

Step 1: Define W^k as the set of active inequality constraints, $W^0 = \varnothing$.

Step 2: Solve the equality constrained quadratic program:

$$\text{Minimize} \quad \frac{1}{2} y^{k\prime} Q y^k + c' y^k$$

$$\text{subject to} \quad e' y^k = 1,$$

$$y_n^k = 0, \quad n \in W^k.$$

Step 3: Let $x^{k+1} = x^k + \alpha^k (y^k - x^k)$ where

$$\alpha^k = \min\left\{\alpha: \alpha \in [0,1],\ x_n^k + \alpha(y_n^k - x_n^k) = 0,\ n = 1,\ldots,N\right\}.$$

Step 4: Update the working set: Let $j \notin W^k$ denote the index of the constraint(s) such that

$$x_j^k + \alpha^k (y_j^k - x_j^k) = 0$$

then

$$W^{k+1} = W^k \cup \{j\}.$$

Step 5: Release a binding constraint: Let $i \in W^k$ denote the index of a constraint such that $\mu_i < 0$, then

$$W^{k+1} = W^k - \{i\}.$$

If $\mu_i > 0$ for all $i \in W^k$ STOP, else GO TO Step 2.

In step 1, the working set, the set of active inequality constraints, is initially defined to be the empty set. It is feasible of course that one or more $x_i^0 = 0$. In this case it would be necessary to define the working set as

$$W^0 = \left\{i: x_i^0 = 0, i = 1,\ldots N\right\}.$$

In step 2, the equality constrained quadratic optimization problem is solved. This problem can be solved using the first-order optimality conditions (Luenberger, 1984) for equality constrained problems by solving the following system of linear equations:

$$\begin{bmatrix} Q & -e & -E^k \\ e' & 0 & 0 \\ (E^k)' & 0 & 0 \end{bmatrix} \begin{pmatrix} y^k \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} -c \\ 1 \\ \underline{0} \end{pmatrix} \tag{15}$$

where E^k is the matrix whose rows are formed by the vectors e_n' that are all zeroes except for a 1 in the n^{th} position for $n \in W^k$. It is not necessary that the solution to (15) yield a solution, y^k that is optimal or feasible to the inequality constrained problem.

In step 3, a line search is conducted from the current candidate solution, x^k, towards y^k. Movement along this direction occurs only until either one of the non-tight inequality constraints becomes tight or until the point y^k is reached (in this case $\alpha^k = 1$). If one or more of the non-tight constraints becomes tight in step 3, they are added as binding constraints in step 4.

In step 5 a constraint that was tight is released if the associated Lagrange multiplier is less than zero. (This condition is based on the first-order optimality conditions for inequality constrained optimization). If none of the Lagrange multipliers are negative and no new constraints are added to the working set, the algorithm is stopped at the optimal solution. If the algorithm is not stopped, steps 2-5 are repeated until an optimal solution is found.

3. 2 A Capacity Allocation Example

This example will utilize the results of the statistical network loading model from Section 2.2 at the intersection of Campbell Avenue and Speedway Boulevard (node 335). Figure 11 shows the layout of the intersection including the proportion of vehicles that turn left, right or proceed though the intersection and the associated phase definition. These values we determined by the City of Tucson's traffic engineering department. Each approach has one turning lane and three through lanes. A saturation flow rate of 1800 vphpl was assumed.

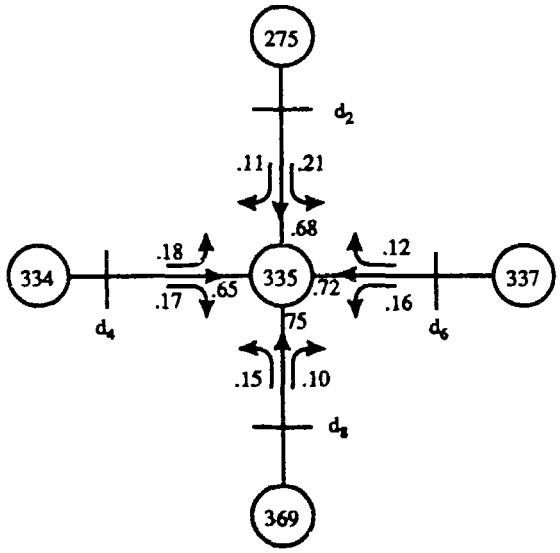


Figure 11. Layout of the Campbell Avenue and Speedway Boulevard intersection.

Table 1 shows the approach volumes for each approach for thirty 15 minute time intervals. These approach volumes are used to derive the movement demands, assuming a 4-phase control, during each time interval. The capacity allocation algorithm is used to find the percent green allocation for each time interval. The results of the capacity allocation algorithm are shown in Table 1.

The results of the capacity allocation algorithm must be carefully interpreted. The numbers in the two tables cannot be directly compared since the approach volumes are related to the traffic volumes using the turning probabilities (see Equation (9)) and the delay is computed using Equation (10). It is also important to note into that the capacity allocation results are to be used for providing estimates and not as the exact signal timing parameters.

Table 1. Approach volumes and percent green allocation from the capacity allocation example. Each approach volume represents the number of vehicles arriving during a 15 minute interval. The percent green allocation is computed to minimize total vehicle delay.

	Approach Volume (per 15 minute period)			
Time	d_2	d_4	d_6	d_8
1	270	416	382	224
2	270	417	384	239
3	266	390	373	231
4	270	433	383	226
5	270	416	382	234
6	266	391	373	212
7	266	391	372	235
8	270	340	384	236
9	270	393	382	239
10	266	382	373	217
11	266	329	371	251
12	270	416	382	241
13	266	447	372	224
14	270	341	379	221
15	270	415	382	235
16	266	392	373	231
17	270	415	383	228
18	270	409	382	236
19	266	392	373	222
20	266	391	373	219
21	270	416	380	247
22	270	414	384	232
23	270	414	384	233
24	270	462	383	212
25	270	418	383	229
26	270	416	383	207
27	270	383	382	217
28	270	468	379	229
29	270	395	384	231
30	266	455	372	221

	Percent Allocation of Green (per Phase)			
Time	Phase 15	Phase 26	Phase 37	Phase 48
1	0.2546	0.2990	0.2104	0.2360
2	0.2533	0.2981	0.2106	0.2379
3	0.2517	0.2945	0.2138	0.2400
4	0.2559	0.3010	0.2086	0.2345
5	0.2536	0.2982	0.2108	0.2375
6	0.2535	0.2963	0.2130	0.2372
7	0.2513	0.2941	0.2140	0.2406
8	0.2473	0.2895	0.2183	0.2449
9	0.2512	0.2950	0.2133	0.2405
10	0.2523	0.2948	0.2141	0.2388
11	0.2447	0.2854	0.2209	0.2490
12	0.2530	0.2975	0.2110	0.2385
13	0.2571	0.3018	0.2077	0.2334
14	0.2485	0.2902	0.2181	0.2432
15	0.2534	0.2979	0.2109	0.2377
16	0.2518	0.2947	0.2136	0.2398
17	0.2542	0.2987	0.2105	0.2366
18	0.2528	0.2971	0.2116	0.2385
19	0.2527	0.2955	0.2133	0.2385
20	0.2529	0.2957	0.2132	0.2382
21	0.2523	0.2967	0.2115	0.2396
22	0.2538	0.2984	0.2107	0.2372
23	0.2537	0.2983	0.2107	0.2373
24	0.2597	0.3058	0.2050	0.2295
25	0.2543	0.2990	0.2103	0.2364
26	0.2562	0.3007	0.2096	0.2335
27	0.2525	0.2958	0.2135	0.2383
28	0.2584	0.3045	0.2054	0.2317
29	0.2523	0.2962	0.2126	0.2389
30	0.2581	0.3030	0.2068	0.2321